



المجلة العراقية للعلوم الاقتصادية
Iraqi Journal For
Economic Sciences



PISSN : 1812-8742

EISSE : 2791-092X

Arcif : 0.375

Advanced Analysis for Evaluating the Efficiency of Multivariate MCUSUM Charts Using Kernel Functions in Air Pollution Monitoring"

تحليل متقدم لتقييم كفاءة لوحات MCUSUM متعددة المخططات
باستعمال دوال كيرنل في مراقبة تلوث الهواء

أم نبا نعيم مهدي

Nabaa Naeem Mahdi

nabaanaeemmahdi@uomustansiriyah.edu.iq

سهاد رافع صالح

Suhad Rafie Saleh

suhad.rafi@uomustansiriyah.edu.iq

كلية الإدارة والاقتصاد / جامعة المستنصرية

Abstract

Air pollution is considered one of the major environmental challenges in Baghdad Province due to the continuous increase in the concentrations of gaseous pollutants resulting from various human activities. Since environmental data are often multivariate and do not follow a normal distribution, the use of nonparametric statistical methods becomes more appropriate for representing their actual characteristics. Among these methods, the Kernel Estimator is a flexible and effective tool for estimating probability distributions without requiring strict distributional assumptions. This study aims to evaluate the efficiency of Multi-Chart MCUSUM control charts based on the Kernel Estimator in monitoring multivariate air pollution data by comparing the performance of three common kernel functions: Gaussian, Laplace, and Epanechnikov. The practical aspect of the study was implemented using R version 4.5.1 due to its advanced capabilities in statistical analysis. The evaluation was based on the ability of the charts to detect small shifts while reducing false alarms, using the proposed Composite Average Run Length (CARL) criterion, which balances chart performance under in-control and out-of-control conditions by integrating the two indicators ARL_0 and ARL_1 into a unified evaluation measure. The results showed that the Gaussian and Laplace functions outperformed the others in terms of speed and sensitivity in detecting small changes, whereas the Epanechnikov function demonstrated lower efficiency in identifying small shifts.

Keywords: Statistical Control Charts, Multi-Chart MCUSUM, Kernel Estimators, Laplace, Epanechnikov, Air Pollutants.

المستخلص

يُعدّ تلوث الهواء من التحديات البيئية الرئيسية في محافظة بغداد نتيجة الارتفاع المستمر في تراكيز الملوثات الغازية الناتجة عن الأنشطة البشرية المختلفة. ونظرًا لأن البيانات البيئية غالبًا ما تكون متعددة المتغيرات ولا تتبع التوزيع الطبيعي، يصبح استعمال الأساليب الإحصائية اللامعلمية أكثر ملاءمة لتمثيل خصائصها الفعلية. من بين هذه الأساليب، يُعدّ مقدّر النواة (Kernel Estimator) أداة مرنة وفعالة لتقدير التوزيعات الاحتمالية دون الحاجة إلى افتراضات توزيعية صارمة. تهدف هذه الدراسة إلى تقييم كفاءة لوحات المجموع المتراكم متعددة المخططات (Multi-Chart MCUSUM) المعتمدة على مقدّر النواة في مراقبة بيانات تلوث الهواء متعددة المتغيرات، من خلال مقارنة أداء ثلاث دوال نواة شائعة هي: Gaussian و Laplace و Epanechnikov. تم تنفيذ الجانب التطبيقي باستعمال برنامج R الإصدار 4.5.1 لما يوفره من إمكانيات متقدمة في التحليل الإحصائي. اعتمد التقييم على قدرة اللوحات في الكشف عن الانحرافات الصغيرة مع تقليل الإندارات الكاذبة، باستعمال معيار المقارنة المركب المقترح (CARL) الذي يوازن بين أداء اللوحة في حالتي السيطرة وخارج السيطرة من خلال دمج مؤشري ARL_1 و ARL_0 ضمن مقياس تقييمي موحد. أظهرت النتائج تفوق دالتي Gaussian و Laplace في سرعة وحساسية الكشف عن التغيرات الطفيفة، في حين أظهرت دالة Epanechnikov كفاءة أقل في رصد الانحرافات الصغيرة.

الكلمات الرئيسية: لوحات السيطرة الإحصائية، MCUSUM متعددة المخططات، مقدر النواة، Laplace, Epanechnikov ملوثات الهواء.

المقدمة

يُعرّف تلوث الهواء بأنه وجود مواد صلبة أو سائلة أو غازية في الغلاف الجوي بتركيز قد تضر بصحة الإنسان والكائنات الحية، أو تعطل النظم البيئية. وتشير التقارير الحديثة إلى أن بغداد تحتل المرتبة الثانية عالميًا بين أكثر المدن تلوثًا بالهواء، وذلك بحسب المنصة السويسرية IQAir المتخصصة في رصد جودة الهواء. ويبرز هذا خطورة الوضع البيئي في المدينة وارتفاع مستويات الملوثات الغازية الناتجة عن الأنشطة البشرية المرتبطة باحتراق الوقود الأحفوري. ويُعد حرق الوقود في محطات توليد الطاقة ووسائل النقل المصدر الرئيس لهذه الانبعاثات، فضلًا عن مصادر أخرى. تُعد السيطرة على النوعية الإحصائية (SPC) من الأدوات الأساسية لضمان استقرار العمليات وتحسين جودة المنتجات والخدمات في المجالات الصناعية والبيئية. ومنذ بداياتها في أوائل القرن العشرين، استُعملت لوحات السيطرة كوسائل فعالة للكشف عن المشكلات وتقليل الهدر وضمان الامتثال للمعايير. ومع التطور التكنولوجي وزيادة تعقيد البيانات وتعدد المتغيرات المترابطة، برزت الحاجة إلى أساليب أكثر تقدمًا، مثل السيطرة على النوعية متعددة المتغيرات (MSPC)، التي تتميز بقدرتها على التعامل مع بيانات مترابطة لا تستطيع الأساليب أحادية المتغير رصدها بدقة. ومن أبرز أدواتها لوحة المجموع المتراكم متعددة المخططات (MCUSUM)، التي تُعد تطورًا جوهريًا بفضل قدرتها على الكشف المبكر عن التغيرات الصغيرة والمفاجئة عبر تحليل التغيرات التراكمية في بيانات العمليات بصورة دقيقة. وأن هذه لوحة لا تقتصر على الكشف عن الانحرافات في العملية، بل تتميز بقدرتها على تحديد المتغير المسبب بخروج المشاهدات عن حدود السيطرة، مما يتيح معالجة السبب الجذري للانحراف وتلافي تكراره مستقبلاً، وهو ما يعزز فاعليتها مقارنة باللوحات التقليدية، فضلًا عن استعمال مقدر الكثافة النووية (KDE) وهو أسلوب إحصائي لا معلمي يُستعمل لتقدير دالة الكثافة الاحتمالية دون افتراض شكل توزيعي محدد، ويعتمد على دوال النواة وعرض الحزمة لتوفير تمثيل مرن وواقعي للبيانات. يُعد من الأدوات الفعالة في مراقبة العمليات متعددة المتغيرات، كونه يتجاوز قيود الأساليب المعلمية التي تفترض ان البيانات تتبع التوزيع الطبيعي.

1. الاستعراض المرجعي: قام الكثير من الباحثين بدراسة لوحات السيطرة على النوعية متعددة المتغيرات ومن بين هذه الدراسات قدّم (Bae & Kvam, 2016) دراسة حول استعمال خرائط

التحكم الإحصائية متعددة المتغيرات اللامعلمية بالاعتماد على مفهوم عمق البيانات كبديل للطرائق التقليدية التي تفترض التوزيع الطبيعي. بينت الدراسة أن مقاييس العمق مثل عمق توكي تساعد في تقليل أبعاد البيانات مع الاحتفاظ بالقدرة على الكشف عن التغيرات الطفيفة والمتوسطة في العملية. كما استعرض الباحثون الأدبيات السابقة وبيّنوا مزايا هذه الأساليب خاصة عندما لا تتبع البيانات التوزيع الطبيعي، بالإضافة إلى توضيح كيفية دمج خرائط EWMA و CUSUM متعددة المتغيرات في هذا الإطار للكشف عن الانحرافات في متوسط العملية. قدّم (Fahad & Jaber, 2020)، دراسة بعنوان استعمال بعض أساليب اللوحات اللامعلمية متعددة المتغيرات في السيطرة على النوعية مع تطبيق عملي، ركزت على تحسين مراقبة الجودة باستعمال أساليب إحصائية حديثة مثل: لوحة، T^2 Hotelling لوحة SN^2 ، طريقة، Bootstrap تقدير كثافة اللب (KDE)، وآلة الدعم المتجهة (SVM)، أظهرت النتائج من خلال البيانات التطبيقية، خصوصاً في معالجة مياه الشرب، تفوق تقنية SVM في تصنيف البيانات بدقة وكفاءة، مؤكدةً فعالية الأساليب اللامعلمية في كشف الانحرافات وتعزيز كفاءة العمليات الإنتاجية والخدمية (Amir, Mohammad 2023)، دراسة بعنوان "مراجعة حول مخططات MCUSUM للكشف عن التغيرات في العمليات مع دراسة مقارنة"، هدفت إلى تقييم شامل لأساليب MCUSUM عبر محاكاة لقياس الانحرافات في المتوسط، والانحراف المعياري، ومعامل الارتباط. أظهرت النتائج أن: طريقة (Crosier 1988) تتفوق في كشف التغيرات الكبيرة في التشتت. وطريقة Pignatiello و (Runger 1990) متميزة في كشف التغيرات الصغيرة والكشف المبكر. وطريقة (Healy 1987) حساسة للتغيرات الإيجابية في معامل الارتباط. كما قدّم (Qiao & Wang, 2024)، دراسة طوّراً فيها لوحة المجموع المتراكم متعددة المتغيرات (MCUSUM) غير المعتمدة على التوزيع بالاعتماد على دوال كيرنل، لمعالجة صعوبات مراقبة البيانات متعددة المتغيرات عند غياب المعرفة المسبقة بتوزيعها وندرة البيانات التاريخية. ركزت الدراسة على: تطوير أدوات فعالة لرصد التغيرات المفاجئة، تحسين تحليل البيانات عالية الأبعاد، التعامل مع محدودية البيانات التاريخية، وأظهرت نتائج المحاكاة والتحليل النظري أن اللوحة المقترحة تقلل ARL1 مع الحفاظ على ARLO مكافئ للمخططات التقليدية (مثل DFMGoF و SSEWMA)، كما قدّم الباحثان مؤشراً جديداً (KCPI) لقياس كفاءة الكشف عن التغيرات.

المبحث الأول

أولاً: مشكلة البحث: تواجه المدن الكبرى مثل بغداد مستويات عالية من تلوث الهواء نتيجة الانبعاثات الناتجة عن حرق الوقود الأحفوري في محطات الطاقة ووسائل النقل، الأمر الذي يشكل تهديداً لصحة الإنسان والنظم البيئية. ونظراً لأن بيانات تلوث الهواء متعددة المتغيرات ولا تتبع بالضرورة التوزيع الطبيعي، فإن الأساليب الإحصائية التقليدية أحادية المتغير أو المعلمية قد تكون غير كافية للكشف المبكر والدقيق عن الانحرافات الطفيفة والمفاجئة في مستويات الملوثات. وتبرز الحاجة إلى تطوير أدوات فعالة لمراقبة هذه البيانات تكون قادرة على التعامل مع البيانات متعددة المتغيرات والمعقدة والكشف المبكر عن التغيرات الصغيرة في متوسط العملية وتقليل الإنذارات الكاذبة الناتجة عن حدود السيطرة التقليدية وتجاوز افتراضات التوزيع الطبيعي باستعمال أساليب لامعلمية مثل مقدر النواة. وبالتالي تتركز مشكلة البحث في تحديد مدى فعالية لوحة المجموع المتراكم متعددة المخططات (MCUSUM) المعتمدة على مقدر النواة في مراقبة تراكيز الملوثات الغازية (SO_2 ، NO_2 ، CO) في بغداد وتحسين دقة الكشف عن التغيرات الصغيرة وتقليل الإنذارات الكاذبة وتحديد المتغير المسؤول

عن خروج المشاهدات عن حدود السيطرة.

ثانياً : هدف البحث: تهدف هذه الدراسة إلى مراقبة تراكيز ملوثات الهواء المتمثلة بثلاث غازات: أول أكسيد الكربون (CO)، وثاني أكسيد النيتروجين (NO₂)، وثاني أكسيد الكبريت (SO₂) في محطة السيديّة ببغداد، من خلال تصميم ومقارنة أداء لوحة المجموع المتراكم متعددة المخططات (Multi-Chart MCUSUM) المعتمدة على دوال مقدر النواة. ويسعى هذا التصميم إلى تحسين قدرة الكشف عن التغيرات الصغيرة والمفاجئة في متوسط العملية متعددة المتغيرات، وتقليل الاعتماد على فرضيات التوزيع الطبيعي أو البيانات التي تتبع توزيعات أخرى، فضلاً عن تعزيز دقة المراقبة الإحصائية للأنظمة البيئية ذات البيانات غير الطبيعية. كما تهدف الدراسة إلى اختبار فعالية المخطط المحسن في تقليل الإنذارات الكاذبة وتحسين حدود التحكم الديناميكية، عبر تطبيقه على بيانات حقيقية لتلوث الهواء في بغداد، وإجراء مقارنات كمية ومعيارية بين لوحات السيطرة، بما يدعم تطوير أنظمة رقابة بيئية أكثر كفاءة وموثوقية.

ثالثاً: هيكلية البحث: يهدف هذا البحث إلى تقييم فعالية لوحة المجموع المتراكم متعددة المخططات (MCUSUM) المعتمدة على مقدر النواة في مراقبة تراكيز ملوثات الهواء، إذ يتناول المبحث الأول الإطار النظري للدراسة ويشمل نبذة عن تلوث الهواء، لوحات السيطرة متعددة المتغيرات، ومقدر النواة مع دواله الثلاث (Gaussian, Laplace, Epanechnikov). أما المبحث الثاني فيركز على الدراسة التطبيقية، التي تعتمد على تحليل بيانات التلوث في محطة السيديّة ببغداد باستعمال MCUSUM متعدد المخططات، ومقارنة أداء دوال النواة المختلفة لتحديد اللوحة الأكثر كفاءة في الكشف عن الانحرافات الصغيرة وتقليل الإنذارات الكاذبة. المبحث الثاني

المحور الاول : الجانب النظري

1. لوحات السيطرة على النوعية متعددة المتغيرات: تعتمد العديد من الدراسات والتطبيقات الإحصائية على لوحات السيطرة لمراقبة متغير، أو خاصية واحدة فقط في العمليات الإنتاجية. ومع ذلك، في الواقع العملي، تتأثر جودة المنتجات بعدة عوامل مترابطة، بدلاً من متغير واحد مستقل، مما يتطلب استعمال تقنيات أكثر تطوراً لمراقبة هذه العوامل بشكل متزامن، لذلك، تم تطوير لوحات السيطرة على النوعية متعددة المتغيرات التي تتيح مراقبة أكثر من متغير واحد وتحليله في آن واحد، مما يساعد على اكتشاف التغيرات الطفيفة التي قد تؤثر في جودة المنتج النهائي، إذ تُعد لوحة المجموع المتراكم متعددة المتغيرات (MCUSUM) من أبرز النماذج المستعملة في هذا المجال، فهي تعتمد على تحليل التغيرات في أكثر من متغير، من خلال حساب إحصاءة المجموع المتراكم، مما يتيح مراقبة العلاقة بين المتغيرات المختلفة، والتأكد من بقائها ضمن الحدود المسموح بها (Hotelling, 1947).

2. لوحة المجموع المتراكم متعددة المخططات باستعمال مقدر النواة: تُعد لوحة المجموع المتراكم متعددة المخططات (Multivariate CUSUM Multi-Chart) من أحدث أساليب المراقبة الإحصائية للعمليات متعددة المتغيرات، حيث صُممت لتجاوز قيود اللوحات التقليدية التي تعتمد على مخطط واحد يستهدف تغيراً محدداً. إن الفكرة الجوهرية لهذه اللوحة هي إنشاء مجموعة من المخططات الفرعية المستقلة، كل منها مهياً لاكتشاف نمط تغير مختلف في بيانات العملية، ثم دمج نتائجها عبر مبدأ (أول إنذار) بحيث يُصدر النظام إشارة عند أول مخطط يتجاوز حد السيطرة، كما تُعد لوحة MCUSUM واحدة من أكثر الأدوات كفاءة في الكشف عن الانحرافات الصغيرة في العمليات متعددة المتغيرات، إذ تعتمد هذه اللوحة على

مبدأ تجميع المعلومات المتراكمة من البيانات لرصد أي انحراف عن المتوسط الطبيعي للعملية وتعد لوحات السيطرة الإحصائية متعددة المخططات (MSPC) أداة مهمة في تحليل العمليات الصناعية والخدمية، إذ تتيح مراقبة العديد من المتغيرات المتداخلة بشكل متزامن، ما يساعد على الكشف المبكر عن التغيرات التي قد تؤثر في جودة العمليات، أو الأداء، وتمثل لوحة السيطرة التراكمية متعددة المتغيرات (MCUSUM) أحد أكثر اللوحات فاعلية للكشف عن التغيرات الصغيرة والمستمرة في العمليات، ومن أهم مميزات سهولة التفسير في التطبيقات العملية، في المجالات الصناعية، أو البيئية مثل: (مراقبة ملوثات الهواء)، إذ تُعد هذه اللوحة أكثر وضوحاً للمختصين، لأنها تُظهر أي المتغيرات تجاوزت حدود السيطرة، أو المتغير المسبب لخروج المشاهدات، مما يساعد في اتخاذ قرارات سريعة ودقيقة (Qiao & Han, 2021).

3. مقدر النواة Kernel Density Estimation – KDE: يُعد مقدر النواة من الأساليب الإحصائية اللامعلمية التي قدّمها Rosenblatt (1952) و Parzen (1962) وطوّرت لاحقاً لتشمل الحالات متعددة المتغيرات. تُستعمل هذه الطريقة لتقدير دالة الكثافة الاحتمالية دون افتراض شكل توزيعي مسبق، مما يجعلها مناسبة عند عدم ملاءمة التوزيعات التقليدية مثل الطبيعي أو ذي الحدين. يهدف مقدر النواة إلى الحصول على تقديرات أكثر دقة لسّمات البيانات الحقيقية، ويتميز بسهولة التطبيق والبرمجة، إضافة إلى استعماله الواسع في التمثيل الرسومي وتحليل البيانات، مما يجعله أداة فعالة في التطبيقات الإحصائية (Parzen, 1962). لنفترض أن X_1, X_2, \dots, X_n تمثل عينة عشوائية من المتغير العشوائي X ، التي تمتلك دالة كثافة احتمالية $f(x)$ غير معروفة. يمكن تقديرها باستعمال مقدر النواة (Kernel Density Estimation)، الذي يعبر عنه بالعلاقة الآتية (Parzen, 1962) (Rosenblatt, 1956):

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k(u) \quad \dots \dots \dots -1$$

$$u = \frac{x-x_i}{h}$$

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right) \quad i = 1, 2, \dots, n \quad \dots \dots \dots -2$$

إذ إن h : عرض النطاق، أو معلمة التمهيد، أو تسمى أيضاً (سعة القيد).
 $K(u)$: دالة النواة التي تعمل كدالة الوزن (Window Function)، وهي دالة حقيقية تقوم بإزالة الانحرافات من المقدار، أي (تعديله). أن دالة النواة المختارة تؤثر في تقدير الكثافة، إذ إن قيمة $\hat{f}(x)$ تكون كبيرة، إذا كان هنالك عدد كبير من المشاهدات تقترب من (x) ، أما إذا كان عدد المشاهدات قليلاً فإن قيمة $\hat{f}(x)$ تكون صغيرة، وفيما يأتي نذكر بعض أنواع دوال (kernel) اللامعلمية الأكثر استعمالاً وهي:

جدول (1) بعض أنواع دوال النواة (kernel) اللامعلمية الأكثر استعمالاً

kernel	K(z)	Support
Gaussian	$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$	$(-\infty, \infty)$
Epanechnikov	$K(u) = \frac{3}{4}(1-u^2)$	$[-1, 1]$
Laplacian	$K(u) = \frac{1}{2} e^{(- u)}$	$[-1, 1]$

إذ إن اختيار دوال النواة (kernel) يعد مهماً وضرورياً للحصول على مقدرات تتقارب من الخواص الإحصائية وإن اختيار عرض الحزمة (Bandwidth) مهم، ولا يقل أهمية عن اختيار دالة النواة الأفضل، لأنه يؤثر في التباين والتحيز (Silverman, 1986), (Turlach, 1993)

1.3 اختيار عرض الحزمة: يُعد اختيار عرض الحزمة (المعروف أيضاً بمعلمة التمهيد Smoothing Parameter) من العوامل الأساسية في التقدير الإحصائي اللامعلمي، إذ يؤثر

بشكل مباشر في جودة تقدير دالة الكثافة الاحتمالية عند استعمال مقدر النواة (Kernel Estimation)، ويؤدي عرض الحزمة دوراً مهماً في تحقيق التوازن بين التحيز (Bias)، والتباين (Variance) في التقدير، ويُعبّر التحيز عن مدى انحراف التقدير عن التوزيع الحقيقي، في حين يُشير التباين إلى درجة التغيير في التقدير مع تغيير العينة. يمكن تلخيص التأثير في التقدير الإحصائي، وهناك عدة طرائق وضعت لتكون قادرة على الحصول على معلمة التمهيد (عرض الحزمة h) الأمثل، ويفضل استعمال طريقة المربعات الصغرى للعبور الشرعي، وذلك لسهولة حسابها، وهيكلتها لأي لوحة سيطرة إحصائية متعددة المتغيرات، لكونها الأكثر استعمالاً من خلال تصغير مربع الخطأ التكاملي (Integrated Squared Error)، وتكمن أهمية هذه العملية في استعمال الطريقة المناسبة التي توازن بين التباين، والانحياز، إذ تقلل نسبة الخطأ الكلي قدر الإمكان. لذا، يُعتمد على عدد من الطرائق الإحصائية لحساب عرض الحزمة الأمثل، ومن أهمها: طريقة المربعات الصغرى للعبور الشرعي (Abdel & Jetter, 2010).

2,3 طريقة المربعات الصغرى للعبور الشرعي: تُعدّ طريقة المربعات الصغرى للعبور الشرعي (LSCV) من أكثر الطرائق التطبيقية شيوعاً لتقدير عرض الحزمة. وقد طُرحت هذه الطريقة لأول مرة من قبل Rudemo (1982) عبر مقياس الخطأ التكاملي المربع (Integrated Squared Error – ISE)، ثم طوّرها Bowman (1984) باعتبارها واحدة من أفضل الأساليب الإحصائية لاختيار القيمة المثلى لعرض الحزمة. ويُعرّف مقياس الخطأ التكاملي بالصيغة الآتية (Abdel & Jetter, 2010):

$$ISE(h) = ISE(\hat{f}_h(x)) = \int_{-\infty}^{\infty} [\hat{f}(x) - f(x)]^2 dx \quad \dots \dots \dots -3$$

إذ إن $f(x)$ لا تعتمد على معلمة التمهيد h ، وأن $\hat{f}(x)$ هو مقدر دالة الكثافة:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)$$

ولكي نحدد قيمة (h) بصورة دقيقة من خلال معرفة فرق القيمة الحقيقية $f(x)$ عن القيمة المقدرة $\hat{f}_h(x)$ وكلما اقتربت كانت قيمة (h) افضل لذا تم استعمال مقياس ISE وبتبسيط معادلة (3) وفتح التربيع نحصل:

$$ISE = \int_{-\infty}^{\infty} [\hat{f}_h(x)]^2 dx - 2 \int_{-\infty}^{\infty} \hat{f}_h(x)f(x)dx + \int_{-\infty}^{\infty} f^2(x)dx \dots \dots -4$$

وبعد التبسيط باستعمال التكامل واستعمال دالة الالتفاف لدالة النواة $k * k(u)$ كالآتي:

$$k * k(u) = \int k(u) * k(u) + R(k)$$

وبعد استعمال دالة الالتفاف والتخلص من التكامل وإجراء بعض العمليات الحسابية تصبح كالآتي:

$$LSCV(h) = \frac{1}{n^2 h^2} \sum_{i=1}^n \sum_{j=1}^n k * k\left(\frac{x_i - x_j}{n}\right) \dots \dots \dots -5$$

وبالتعويض نحصل على:

$$LSCV(h) = \frac{1}{n^2 h^2} \sum_{i=1}^n \sum_{j=1}^n k * k\left(\frac{x_i - x_j}{n}\right) - \frac{2}{n} \sum_{i=1}^n \hat{f}_{-i}(x) \dots \dots \dots -6$$

$$h_{LSCV} = \operatorname{argmin}_h LSCV(h) \dots \dots \dots (7)$$

إذ إن:

1. $f(x)$: الدالة الحقيقية للكثافة الاحتمالية.
2. $\hat{f}_h(x)$: مقدر دالة الكثافة باستعمال النواة مع معامل التمهيد h .
3. h : معامل التمهيد.

4. n: حجم العينة (عدد المشاهدات).

5. x_i : المشاهدات رقم i .

6. $K(\cdot)$: دالة النواة وهي دالة متماثلة.

7. K^*K : التفاف دالة النواة.

8. $\hat{f}_{-i}(x_i)$: المقدر بدون العنصر i .

9. $K(0)$: قيمة دالة النواة عند الصفر.

10. $\arg\min_n LSCV(h)$: قيمة h التي تحقق اقل قيمة لـ $LSCV(h)$ وهي تمثل معامل التمهيد الأمثل (Silverman, 1986).

4. آلية عمل لوحة المجموع المتراكم متعددة المتغيرات MCUSUM باستعمال مقدر كيرنل:

تعد طريقة حديثة لكونها تعتمد على مقدر النواة لبناء لوحات سيطرة متعددة المتغيرات غير معتمدة على التوزيع. وبناء على المشاهدات السابقة m ، يقدر متجه المتوسط قبل التغير ($\hat{\mu}_0$) في حين متوسط ما بعد التغير (μ_1) يبقى مجهولاً عادة في بيئة المراقبة اللحظية، ولمواجهة هذه المشكلة نفترض وجود مجال ممكن لقيم μ_1 يرمز له بـ D ، مستندين إلى البيانات السابقة، أو الخلفية النظرية للبيانات، ومن هذا المجال يتم اختيار C متجهات مرجعية ممكنة لما بعد التغير ترمز بـ μ_1^k حيث $1 \leq k \leq C$ يهدف إلى تغطية احتمالات مختلفة من التحولات المحتملة في البيانات، ويدمج هذه المتجهات المرجعية مع دالة النواة المناسبة لنشئ مخططات فرعية متعددة لاكتشاف التغيرات.

إذ يعرف المخطط الكلي T_{MC} وفق الصيغة الآتية:

$$T_{MC} = \min_{1 \leq k \leq C} T_k \quad \dots \dots \dots -8$$

ويعرف المخطط الفرعي T_k وفق الصيغة الآتية:

$$T_k = \min \left\{ 1 \leq n \leq N: \max_{1 \leq j \leq n} \sum_{i=j}^n \ln \frac{k(X_i, \mu_1^k)}{k(X_i, \hat{\mu}_0)} > d_n^k \right\} \dots \dots \dots -9$$

إذ إن:

1. T_{MC} : إحصاءة لوحة MCUSUM متعددة المتغيرات.
2. T_k : قيمة اختبار فردي في اللوحة.
3. K : مؤشر للمخطط الفرعي، إذ $K=1,2,\dots,C$ و C هو عدد المخططات الفرعية.
4. N : عدد المشاهدات الكلية في العملية.
5. \ln : اللوغاريتم الطبيعي، يستعمل لحساب نسبة احتمالية التغير.
6. $K(\cdot, \cdot)$: دالة النواة (Kernel function)، مثل: Gaussian، أو Laplace، لحساب درجة التشابه بين المشاهدة، والمتجه المرجعي.
7. $K(X_i, \mu)$: دالة النواة التي تربط المشاهدة X_i بالمتوسط μ .
8. μ_0 : المتوسط قبل التغير (في حالة تحت السيطرة).
9. μ_k : المتجه المرجعي بعد التغير للمخطط الفرعي K ، ويمثل نوع تغير محدد متوقع.
10. d_n^k : حد السيطرة للمخطط الفرعي T_k عند المشاهدة رقم n (Qiao & Han, 2021) (Qiao & Wang, 2024).

1.4 دوال النواة المستعملة في لوحة المجموع المتراكم متعددة المخططات MCUSUM:

تعتمد لوحة المجموع المتراكم متعددة المخططات MCUSUM على تعريف دالة النواة المناسبة، التي تُمكننا من بناء إحصائيات قوية وفعالة، تُستعمل في اكتشاف التغيرات داخل

البيانات المتعددة المتغيرات. بناءً على هذه الدالة، تُطور لوحة المجموع المتراكم متعددة المخططات المستندة إلى دوال النواة، مع تحليل الخصائص النظرية، لضمان استقرار أدائه ودقته في التنبؤ بالتغيرات. ويشمل ذلك دراسة مدى حساسية المخطط للتغيرات الطفيفة، بالإضافة إلى تحليل خواصه الرياضية، لضمان موثوقيته في التطبيقات العملية. وقد جرى الاعتماد في هذه الدراسة على ثلاث دوال وهي: نواة Gaussian، ونواة Laplace، ونواة Epanechnikov، لخصائصهما المفيدة في تحليل البيانات، إذ تمتاز نواة Gaussian بقدرتها على التعامل مع التغيرات التدريجية، في حين تُستعمل نواة Laplace و Epanechnikov، لاكتشاف التغيرات السريعة، والمفاجئة (Qiao & Wang, 2024).

2.4 دالة النواة Gaussian: وتحسب وفق الصيغة الآتية :

$$K_{G,h}(x,y) = \exp\left(-\frac{\|x-y\|^2}{2h^2}\right) \dots \dots \dots - 10$$

اذ ان :

1. $\|x - y\|^2$: مربع المسافة الإقليدية بين النقطتين x, y .

2. h : عرض الحزمة (Bandwidth).

3.4 دالة النواة Laplace :

وتحسب وفق الصيغة الآتية (Montgomery, 2009) (Qiao & Wang, 2024):

$$K_{L,h}(x,y) = \exp\left(-\frac{\|x-y\|}{h}\right) \dots \dots \dots - 11$$

4.4 دالة النواة Epanechnikov: دالة النواة Epanechnikov تكون حسب الصيغة الآتية:

$$K_{E,h}(x,y) = \left(1 - \frac{\|x-y\|^2}{h^2}\right) \dots \dots \dots - 12$$

5 . حساب حد التحكم الديناميكي $\{d_n^k\}$: هو مفهوم إحصائي أساسي في تصميم لوحة المجموع المتراكم متعددة المتغيرات (MCUSUM)، ويُستعمل لتحديد النقطة الزمنية التي يجري عندها إصدار إنذار بوجود تغير في العملية الإنتاجية، أو الخدمية. يتميز هذا الحد بمرونته الديناميكية، إذ يتغير مع الزمن، وعدد المشاهدات، مما يجعله أكثر ملاءمة للعمليات ذات الطبيعة المتغيرة، أي عند بداية العملية (عند القيم الصغيرة لـ n)، يتم تعيين حد تحكم أولي لتوفير حساسية للكشف عن التغيرات المبكرة، ومع زيادة عدد المشاهدات n ، واستمرار مراقبة العملية، يُعدل هذا الحد ديناميكياً، بناءً على أداء العملية، ويُحسب هذا الحد بطريقة غير مباشرة، إذ لا يعتمد على صيغة مغلقة كما في بعض المخططات التقليدية، وإنما يُحدد عبر المعايير الإحصائية (Statistical Calibration)، ولأن هذه المخططات لا يمكن تمثيلها بسلسلة احتمالية ثابتة، فقد جرى الاعتماد في حسابه على محاكاة مونت كارلو، لتجريب قيم مختلفة لحد السيطرة حتى يتم الوصول إلى قيمة تحقق متوسط طول التشغيل (ARL₀) تحت السيطرة عند المستوى المطلوب. وبذلك يُستنتج d_n^k عملياً من خلال التجربة الإحصائية حتى يوازن بين سرعة الاكتشاف، وتقليل الإنذارات الكاذبة، وباختصار، فإن d_n^k في جميع اللوحات أعلاه لا يُحسب مباشرة، وإنما يُحدد عن طريق معايير إحصائية دقيقة تأخذ بعين الاعتبار حجم الانحراف المراد اكتشافه، والترابط بين المتغيرات، ومتوسط طول التشغيل المستهدف. وهذا النهج يضمن أن اللوحتين توفران توازناً عملياً بين الكشف المبكر عن التغيرات في متوسط العملية، وتقليل احتمالية إصدار إنذارات خاطئة (Hamed et al., 2016) (Qiao & Han, 2021) (Qiao & Wang, 2024). والفرضية التي جرى اختبارها تكون:

العملية تحت السيطرة $H_0: \mu = \mu_0$. العملية خارج حدود السيطرة $H_1: \mu \neq \mu_0$

6 . معيار المقارنة المركب المقترح Composite Average Run Length Criterion

(CARL): يُستعمل هذا المعيار لقياس كفاءة لوحات السيطرة النوعية، ولمعرفة أيها أفضل أداءً، إذ كلما كانت قيمته أصغر دلّ ذلك على أنّ اللوحة أكثر دقة في الاكتشاف. ويعتمد هذا المعيار على دمج معيارين أساسيين هما: متوسط طول التشغيل في حالة السيطرة ARL_0 ، ومتوسط طول التشغيل في حالة الانحراف أو خارج السيطرة ARL_1 ، وبذلك يمكن إجراء مقارنة بين اللوحات المختلفة، وتحديد أيها الأفضل، لقد جرى اعتماد هذا المعيار المركب المقترح لان في أغلب الدراسات السابقة المتعلقة بلوحات السيطرة النوعية، يُعتمد متوسط طول التشغيل في حالة الانحراف، أو خارج حدود السيطرة ARL_1 كمعيار أساس للمقارنة بين أداء اللوحات، وذلك من خلال تثبيت قيمة عند متوسط مستهدف $target ARL_0$ واحد لجميع اللوحات، ثم تحديد اللوحة الأكفأ، استناداً إلى اللوحة التي تحقق أصغر قيمة لـ ARL_1 . غير أن هذا الأسلوب يُعدّ غير دقيق إحصائياً، إذ لا يمكن إهمال دور ARL_0 أو افتراض ثباته عند قيمة مستهدفة واحدة لجميع اللوحات، من دون النظر إلى الانحرافات الناتجة عنه. ولغرض معالجة هذه المشكلة، جرى في هذه الرسالة اقتراح معيار مركب، يجمع بين الدقة في تحقيق ARL_0 ، والقوة في سرعة اكتشاف الإنذارات الكاذبة المتمثلة بـ ARL_1 . إذ يقوم هذا المعيار على الموازنة بين أصغر قيمة لـ ARL_1 التي تعكس سرعة اكتشاف الانحرافات، ومدى قرب القيمة المقدّرة لـ ARL_0 من المتوسط المستهدف. ويُعرّف هذا المعيار بالصيغة الآتية:

$$CARL = \omega_0 * ARL_{0-der} + \omega_1 * ARL_{1-rel} \dots \dots \dots -13$$

إذ إن :

1. ω_0 : وزن يستعمل لقياس دقة ARL_0 .
2. ω_1 : وزن يستعمل لقياس سرعة الاكتشاف المرتبطة بـ ARL_1 ، أي يعطي وزناً لسرعة الاكتشاف.
3. ARL_{0-der} : انحراف ARL_0 المحسوبة عن القيمة المستهدفة ($ARL_{0-target}$) عند (α)
4. ARL_{1-rel} : النسبة لمعدل سرعة الاكتشاف لكل لوحة .
- 1 - لحساب ARL_{0-der} : يحسب وفق الصيغة الآتية :

$$ARL_{0-der} = \frac{|ARL_{0-target} - ARL_{0-المحسوبة}|}{ARL_{0-target}} \dots \dots \dots -14$$

1. إذ إن هذه الصيغة تقيس مدى انحراف القيمة المحسوبة لـ ARL_0 (أي متوسط طول التشغيل تحت السيطرة) عن القيمة المستهدفة التي يجري تحديدها مسبقاً $ARL_{0-target}$.
2. إذا كانت القيمة المحسوبة مطابقة تماماً للقيمة المستهدفة ، فإن الفرق في البسط يساوي صفراً، أي $ARL_{0-der} = 0$.
3. كلما ابتعدت القيمة المحسوبة عن الهدف ازدادت قيمة الانحراف النسبي .
4. الهدف أن تكون هذه القيمة قريبة جداً من الصفر، مما يعني أن اللوحة تضبط مستوى الإنذارات الكاذبة بدقة، ولا تطلق إشارات، أو إنذارات أقل أو أكثر مما هو متوقع .
- 2 - لحساب ARL_{1-rel} : من خلال الصيغة الآتية :

$$ARL_{1-rel} = \frac{ARL_1}{\min(ARL_1)} \dots \dots \dots -15$$

1. هذا القانون يقيس الكفاءة النسبية للسرعة في اكتشاف الانحرافات .
2. ARL_1 : يمثل متوسط طول التشغيل عندما تكون العملية خارج حدود السيطرة .
3. $\min(ARL_1)$: هي أصغر قيمة لـ ARL_1 بين جميع اللوحات قيد المقارنة .

4. عندما تكون قيمة ARL_{1-rel} قريبة من 1 فهو يعني أن اللوحة قريبة جداً من أسرع لوحة في الاكتشاف .

المبحث الثالث/المحور الثاني: الجانب التطبيقي

1. المتغيرات المساهمة في تلوث الهواء في بغداد- محطة السيديية

1.1 غاز أحادي أكسيد الكربون CO: يتميز هذا الغاز بأنه ليس له لون، أو طعم، أو رائحة، ويوجد في الهواء الطبيعي غير الملوث عند تركيز لا يتجاوز 0,1 جزء من المليون. ينتج هذا الغاز من احتراق غير كامل للوقود، وخصوصاً في وسائل النقل، وذلك لعدم وجود الأوكسجين الكافي للاحتراق، ويُعد هذا الغاز من أشد الغازات الملوثة للهواء سمية، بسبب قدرته على الاتحاد مع هيموغلوبين الدم بدلا من الأوكسجين، مكوناً كربوكسي-هيموغلوبين، إذ إن زيادة تركيز CO في الجو عن 9 أجزاء من المليون تسبب ضعفاً في القوة العضلية، وضعف السمع، ونقصاً في الرؤية، وغثياناً، وانخفاضاً في الحرارة، وإعاقة في الأمعاء، وقد تؤدي إلى الوفاة خلال ساعات (Central Statistical Organization, 2023).

2.1 أكاسيد النيتروجين (NO_x) Oxides of nitrogen: يقصد بأكاسيد النيتروجين مركبات النيتروجين الغازية التي تتكون عند اتحاد النيتروجين، والأوكسجين تحت درجات حرارة، عالية كما هو الحال عند احتراق البنزين في المركبات، كالسيارات، وغيرها . ومن أهم هذه المركبات أول أكسيد النيتروجين NO، وثنائي أكسيد النيتروجين NO₂، وهذه الغازات سامة، إذ تكون حامض النتريك، وقد تؤدي إلى الموت خلال نصف ساعه إذا وصلت نسبتها في الهواء إلى (0.07%) وتتحد هذه الأكاسيد أيضاً مع هيموغلوبين الدم، وتمنع نقل الأوكسجين إلى الخلايا (Central Statistical Organization, 2023).

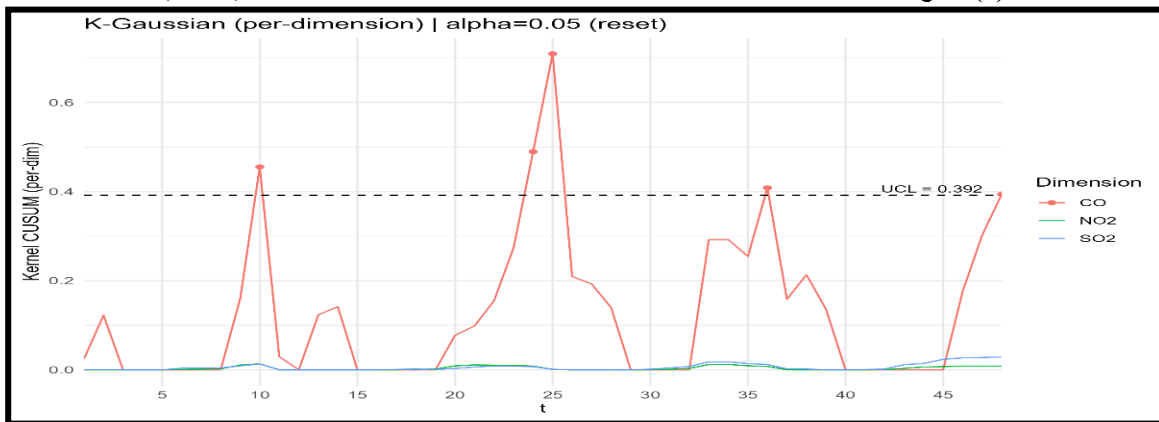
3.1 أكاسيد الكبريت (SO_x) Sulphur oxide : وتشمل ثنائي أكسيد الكبريت SO₂، وثالث أكسيد الكبريت SO₃، إذ يمتاز غاز ثنائي أكسيد الكبريت SO₂ عديم اللون وذو رائحة حادة يؤدي إلى السعال، وينتج القسم الأعظم منه بسبب احتراق الوقود الأحفوري، إذ يحتوي الوقود الأحفوري على نسب مختلفة من الكبريت الذي يخرج إلى الغلاف الجوي على شكل ثنائي أكسيد الكبريت عند احتراقه، ويتفاعل غاز ثنائي أكسيد الكبريت مع أوكسجين الهواء، وبخار الماء، لينتج حامض الكبريتيك . وتتعرض المباني وخاصة القديمة الأثرية إلى تأثير حامض الكبريتيك الذي يؤدي إلى تآكل أو تجوية هذه المباني (Central Statistical Organization, 2023) (Al-Dabbagh , 2022)

الشكل (1): انبعاثات الغازات الملوثة من مداخل المصانع في بغداد



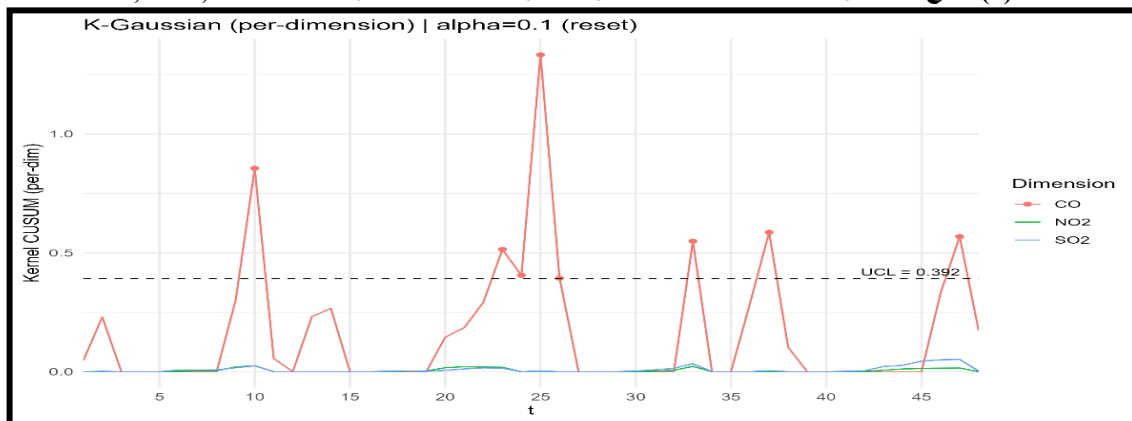
2. وصف البيانات وتحليلها: تعد محافظة بغداد المركز الإداري، والسياسي، والاقتصادي لجمهورية العراق، وهي المدينة الأكثر اكتظاظًا بالسكان، إذ تحتضن ما يزيد على 8 ملايين نسمة ضمن مساحة حضرية محدودة نسبيًا، وهذا التوسع السكاني، والعمراني السريع، إضافة إلى الاعتماد الكبير على وسائل النقل التقليدية، والمولدات الكهربائية، جعل من بغداد إحدى أكثر المدن العراقية تعرضًا لتلوث الهواء، لا سيما في المناطق الحضرية ذات الكثافة السكانية العالية، وتشير تقارير وزارة البيئة العراقية، ومنظمة الصحة العالمية إلى أن مستويات تلوث الهواء في بغداد تتجاوز في كثير من الأوقات الحدود المسموح بها دوليًا، مما ينعكس سلبيًا على الصحة العامة، وسلامة البيئة، ويستدعي مزيدًا من الدراسات التحليلية لفهم اتجاهات التلوث، ومصادره بشكل دقيق. استنادًا إلى ذلك، جرى في هذه الدراسة تحليل بيانات الرصد البيئي الشهري الصادرة عن محطة السيدية في بغداد، وهي إحدى المحطات المعتمدة لرصد جودة الهواء، إذ امتدت البيانات من شهر كانون الثاني (2018 /1)، إلى كانون الأول (2022 /12)، وبواقع 60 قراءة شهرية لكل ملوث من الملوثات الثلاثة الآتية: CO غاز أحادي أوكسيد الكربون، و NO₂ غاز ثنائي أوكسيد النتروجين، و SO₂ غاز ثنائي أوكسيد الكبريت، التي اختيرت بناءً على أهميتها الصحية، وشيوعها في البيئة الحضرية، وتم تحليل البيانات واستخراج النتائج باستعمال برنامج مكتوب باللغة R الاصدار 4.5.1 لما يوفره من إمكانيات متقدمة في التحليل الإحصائي وكانت النتائج كالآتي:

شكل (2) نتائج دالة النواة Gaussian، عند مستوى معنوية $\alpha=0.05$ ، للمتغيرات CO, NO₂, SO₂.



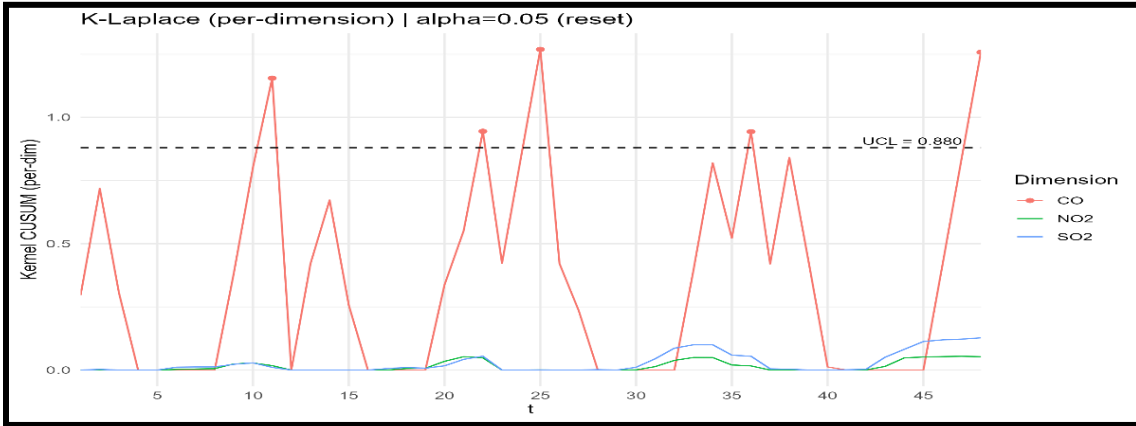
في الشكل (2) عندما يكون مستوى المعنوية $\alpha=0.05$ وعرض الحزمة ($h=0.122$)، وللملوثات الثلاثة (CO, NO₂, SO₂)، فإن العينات (10,24,25,36,48)، تقع خارج حدود السيطرة، مع حد السيطرة الأعلى $d_n^k = 0.392$.

شكل (3) نتائج دالة النواة Gaussian، عند مستوى معنوية $\alpha=0.1$ ، للمتغيرات الثلاثة CO, NO₂, SO₂.



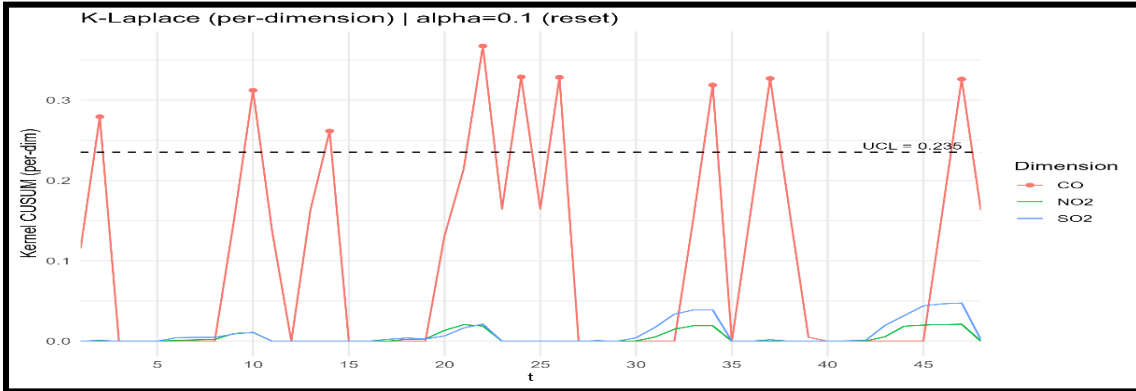
في الشكل (3) عندما يكون مستوى المعنوية $\alpha=0.1$ وعرض الحزمة ($h=0.089$)، وللملوثات الثلاثة (CO, NO_2, SO_2)، فإن العينات (10,23,24,25,26,33,37,47)، تقع خارج حدود السيطرة، مع حد السيطرة الأعلى $d_n^k = 0.392$.

شكل (4) نتائج دالة النواة Laplace، عند مستوى معنوية $\alpha=0.05$ ، وللمتغيرات الثلاثة CO, NO_2, SO_2



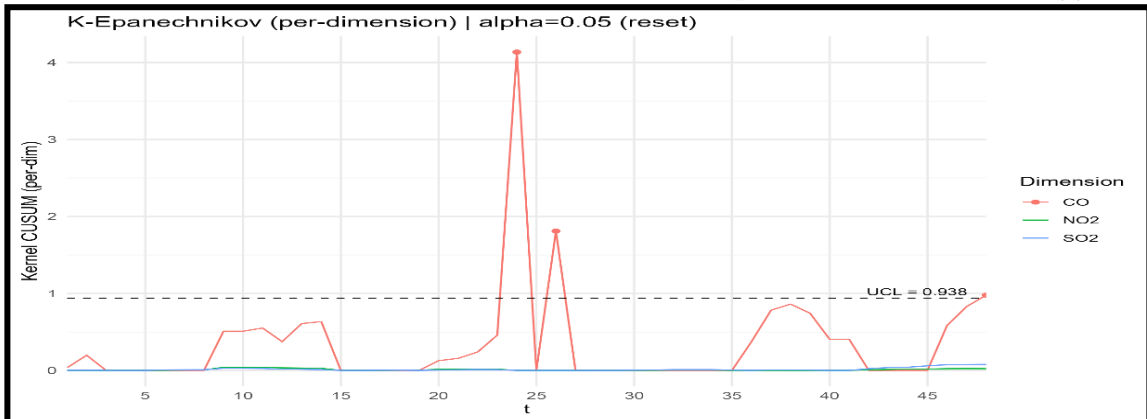
في الشكل (4) عندما يكون مستوى المعنوية $\alpha=0.05$ وعرض الحزمة ($h=0.0560$)، وللملوثات الثلاثة (CO, NO_2, SO_2)، فإن العينات (11,22,25,36,48)، تقع خارج حدود السيطرة، مع حد السيطرة الأعلى $d_n^k = 0.880$.

شكل (5) نتائج دالة النواة Laplace، عند مستوى معنوية $\alpha=0.1$ ، للمتغيرات الثلاثة CO, NO_2, SO_2 .

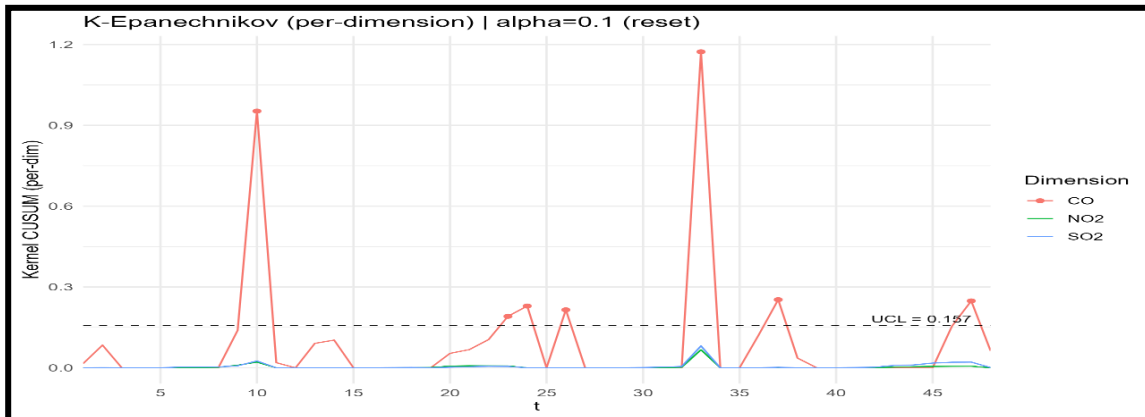


في الشكل (5) عندما يكون مستوى المعنوية $\alpha=0.1$ وعرض الحزمة ($h=0.144$)، وللملوثات الثلاثة (CO, NO_2, SO_2)، فإن العينات (2,10,14,22,24,26,34,36,47)، تقع خارج حدود السيطرة، مع حد السيطرة الأعلى $d_n^k = 0.235$.

شكل (6) نتائج دالة النواة Epanechnikov، عند مستوى معنوية $\alpha=0.05$ ، للمتغيرات الثلاثة CO, NO_2, SO_2 .



في الشكل (6) عندما يكون مستوى المعنوية $\alpha=0.05$ ، وعرض الحزمة ($h=0.148$)، وللملوثة الثلاثة (CO, NO_2, SO_2)، فإن العينات (24,26,48)، تقع خارج حدود السيطرة، مع حد السيطرة الأعلى $d_n^k = 0.938$. شكل (7) نتائج دالة النواة Epanechnikov، وعند مستوى معنوية $\alpha=0.1$ ، للمتغيرات الثلاثة CO, NO_2, SO_2 .



في الشكل (7) عندما يكون مستوى المعنوية $\alpha=0.1$ ، وعرض الحزمة ($h=0.216$)، وللملوثة الثلاثة (CO, NO_2, SO_2)، فإن العينات (10,23,24,26,33,37,47)، تقع خارج حدود السيطرة، مع حد السيطرة الأعلى $d_n^k = 0.157$. ويتبين من نتائج التطبيق العملي لمقدر النواة على لوحة المجموع المتراكم متعددة المخططات (MCUSUM)، أن أداء دوال النواة يختلف باختلاف مستوى المعنوية المعتمد، مع تسجيل فروقات واضحة في عدد المشاهدات الخارجة عن حد السيطرة الأعلى (d_n^k). دالة Gaussian: عند مستوى معنوية 0.05 سجلت خروج 5 مشاهدات عن حد السيطرة الأعلى، بينما عند مستوى معنوية 0.1 ارتفع العدد إلى 8 مشاهدات. دالة Laplace: عند مستوى معنوية 0.05 حققت نتائج مشابهة لدالة Gaussian بخروج 5 مشاهدات، إلا أنها عند مستوى معنوية 0.1 تفوقت قليلاً بتسجيل 9 مشاهدات خارج السيطرة. دالة Epanechnikov: كانت الأقل كفاءة من حيث التحسس للتغيرات، إذ سجلت 3 مشاهدات عند مستوى معنوية 0.05 وارتفعت إلى 7 مشاهدات عند مستوى معنوية 0.1.

المبحث الرابع / الاستنتاجات والتوصيات.

- 1. الاستنتاجات:** نستنتج من التطبيق العملي لمقدر النواة على لوحة المجموع المتراكم متعددة المخططات (MCUSUM)، أن أداء دوال النواة يختلف باختلاف مستوى المعنوية المعتمد، مع تسجيل فروقات واضحة في عدد المشاهدات الخارجة عن حد السيطرة الأعلى (d_n^k) وكما موضح:
 1. دالة Gaussian: عند مستوى معنوية 0.05 سجلت خروج 5 مشاهدات عن حد السيطرة الأعلى، بينما عند مستوى معنوية 0.1 ارتفع العدد إلى 8 مشاهدات، أما دالة Laplace: عند مستوى معنوية 0.05 حققت نتائج مشابهة لدالة Gaussian بخروج 5 مشاهدات، إلا أنها عند مستوى معنوية 0.1 تفوقت قليلاً بتسجيل 9 مشاهدات خارج السيطرة، أما دالة Epanechnikov: كانت الأقل كفاءة من حيث التحسس للتغيرات، إذ سجلت 3 مشاهدات عند مستوى معنوية 0.05 وارتفعت إلى 7 مشاهدات عند مستوى معنوية 0.1.
 2. أظهرت نتائج الجانب التطبيقي أن متغير أول أكسيد الكربون (CO)، كان الأكثر تلوثاً مقارنة ببقية الملوثات (SO_2, NO_2)، إذ سجل أعلى مستويات انحراف عن حدود السيطرة في البيانات الحقيقية، مما يجعله العامل الرئيس في خروج العملية البيئية عن السيطرة ويؤكد الحاجة إلى مراقبته بشكل مكثف. في المقابل، بينت النتائج أن الملوثات الأخرى (SO_2, NO_2)، كانت ضمن حدود السيطرة، وهو ما يشير إلى استقرارها النسبي وعدم تسببها بخروج المشاهدات عن السيطرة خلال فترة الدراسة.

2. التوصيات

1. نوصي وزارة البيئة بأعتماد لوحات المجموع المتراكم متعددة المخططات (MCUSUM)، باستعمال مقدر النواة لرصد مستويات تلوث الهواء في بغداد، مع التركيز على الملوثات الرئيسية (SO_2 ، NO_2 ، CO)، لما لها من تأثيرات مباشرة في صحة الإنسان، إذ أثبتت هذه اللوحات قدرتها على الكشف المبكر عن الانحرافات. كما يُقترح توسيع نطاق المراقبة ليشمل ملوثات أخرى، مثل: (O_3 ، CH_4 ، NO_x)، والجسيمات العالقة (PM)، فضلاً عن تطبيق هذه المنهجية في محطات الرصد البيئي في باقي المحافظات، بما يعطي صورة أشمل، وأكثر دقة لواقع التلوث البيئي في العراق .
2. ضرورة متابعة الوزارة لمستويات أول أكسيد الكربون CO نظراً لارتفاعه الملحوظ في البيانات الحقيقية، وعده من أخطر الملوثات على الصحة العامة، مما يستدعي زيادة حملات القياس المستمرة، ووضع خطط إنذار مبكر عند تجاوز الحدود وفقاً للمعايير العالمية (WHO, EPA).

المصادر Reference

1. Abdel, R. and Jetter, J. (2010). "A Simulation Study For The Bandwidth Selection In The Kernel Density Estimation Based On The Exact and the Asymptotic MISE", Pak. J. Statist., Vol. 26, No. (1), pp. 239–265.
2. Al-Dabbagh, D. F. (2022). Environmental pollution (air, water, soil, and radioactive pollution) in Iraq in 2022. Iraqi Council of Representatives, Directorate of Research and Studies / Research Department.
3. Bae, S.J., Do, G., & Kvam, P. (2016). "On Data Depth and the Application of Nonparametric Multivariate Statistical Process Control Charts". University of Richmond, Math and Computer Science Faculty Publications, No. 168 .
4. Central Statistical Organization. (2023). Environmental statistics of Iraq: Air pollutants for 2022. Ministry of Planning, Republic of Iraq. Available at: <http://www.cosit.gov.iq>.
5. Enad, F. H., & Jaber, A. G. (2020). The use of some nonparametric multivariate control chart methods in quality control with an applied application. Journal of Economic and Administrative Sciences, 26(124), 443–459.
6. Fallahnezhad, M.S., & Ghalichehbab, A. (2023). "A review on the MCUSUM charts in detecting the shifts of the process with comparison study". International Journal of Innovation in Engineering, Vol. 3, No. 2, pp. 30-38.
7. Hamed, M. S., Mansour, M. M., & Abd Elrazik, E. M. (2016). MCUSUM control chart procedure: Monitoring the process mean with application. Journal of Statistics: Advances in Theory and Applications, 16(1), 105–132 .
8. Hotelling, H. Multivariate quality control—Illustrated by the air testing of sample bombsights. In *Techniques of Statistical Analysis*; Eisenhart, C., Hastay, M.W., Wallis, W.A., Eds.; McGraw-Hill: New York, NY, USA, 1947; pp. 111–184.
9. Montgomery, D. C (2009) Introduction to statistical quality control ". John wiley sons.
10. Parzen, E. (1962). 'On estimation of a probability density function and mode', Ann. Stat. Vol.33, PP1065–1076.
11. Qiao, L.; Han, D (2021) CUSUM multi-chart for detecting unknown abrupt changes under finite measure space for network observation sequences". Statistics, 2021, 55,–513.
12. Qiao, L., & Wang, B. (2024). "Kernel-Based Multivariate Nonparametric CUSUM Multi-Chart for Detection of Abrupt Changes. Mathematics, 12(1473) .
13. Rosenblatt, M. (1956). 'Remarks on some nonparametric estimates of a density function', Ann. Math. Stat. Vol. 27, NO. 3, PP.832–837.
14. Silverman, B.W. (1986). "Density estimation for statistics and data analysis", Chapman and Hall, Londo.
15. Turlach, B. A. (1993). "Bandwidth Selection in Kernel Density Estimation: A Review". C.O.R.E. and InstItut de Statistique. Université Catholique de Louvain.